UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/808,326	03/25/2004	Shioupyn Shen	0026-0072	5423

44989 7590 02/03/2009

HARRITY & HARRITY, LLP
11350 Random Hills Road
SUITE 600
FAIRFAX, VA 22030

EXAMINER	
SMITH, GARRETT A	

ART UNIT	PAPER NUMBER
2168	

MAIL DATE	DELIVERY MODE
02/03/2009	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)	
	10/808,326	SHEN, SHIOUPYN	
	Examiner	Art Unit	
	Garrett Smith	2168	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *19 November 2008*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1,3-9,11-15,17-26 and 28-36* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1, 3-9, 11-15, 17-26, and 28-36* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

1. This Office Action is regarding Applicant's response filed 19 November 2008 to a prior Office Action. Claims 1, 3 – 9, 11 – 15, 17 – 26 and 28 – 36 are pending. Claims 1, 3 – 9, 11 – 15, 17 – 20, 22 – 26 and 28 – 34 are amended. Claims 2, 10, 16 and 27 are canceled. Claims 35 and 36 are new.

2. This Office Action is the **Fourth Action, Non-Final Rejection**.

Continued Examination under 37 CFR 1.114

3. A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed 19 November 2008 in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on 20 October 2008 has been entered.

Response to Arguments

35 USC § 102(a)

4. Applicant's arguments (page 12 – 17) and amendments, filed 20 October 2008, regarding the rejection under 35 USC § 102(a) of claims 13 and 17 have been fully considered but they are not persuasive.

The Examiner notes that Applicant's arguments center around the amendments to the claims regarding "checksum values". The differences between the previous set of

claims and the newly amended claims is the use of "checksums" rather then "blocks". The Examiner submits that the use of a "checksum" to generate a fingerprint is very old. The definition of a checksum is:

> A computed value which depends on the contents of a block of data and which is transmitted or stored along with the data in order to detect corruption of the data. The receiving system recomputes the checksum based upon the received data and compares this value with the one sent with the data. If the two values are the same, the receiver has some confidence that the data was received correctly. The checksum may be 8 bits (modulo 256 sum), 16, 32, or some other size. It is computed by summing the bytes or words of the data block ignoring overflow. The checksum may be negated so that the total of the data words plus the checksum is zero. Internet packets use a 32-bit checksum. See also digital signature, cyclic redundancy check. (1996-03-01) ("checksum." The Free On-line Dictionary of Computing. Denis Howe. 01 Feb. 2009. <Dictionary.com http://dictionary.reference.com/browse/checksum>).

The Examiner submits that the definition of checksum renders obvious a number of the limitations of the claims by itself. The Examiner further submits that Schleimer can be easily modified to use a “checksum” as shown in the above definition rather then a text block. However, the Examiner also notes that Schleimer does also discuss “hashing” of the various blocks. “Hashing” can also be reasonable considered to be generation of a checksum value. Applicant’s specification discusses how checksum functions are “hash functions” (see [0037]) and that such checksum functions are well known in the art.

The Examiner further notes that there is no requirement that checksum values generated have to have any relation to the overlapping blocks. The checksums are generated “for” the blocks but not “based on” the blocks. The Examiner submits the amended phrase “where a number of times a particular bit is flipped is based on a number of checksum values in subset that corresponds to the particular value

addressed to the particular bit" is intended use. The phrase does not appear to modify anything in the claim.

For these reasons, the rejection under 35 USC § 102(a) of claims 13 and 17 is **maintained**.

35 USC § 103(a)

5. Applicant's arguments (page 17 – 35) and amendments, filed 20 October 2008, regarding the rejection under 35 USC § 103(a) of claims 1, 3 – 9, 11 – 15, 17 – 26 and 28 – 34 have been fully considered but they are not persuasive. See arguments provided in the 102(a) section above. For these reasons, the rejection under 35 USC § 103(a) of claims 103(a) of claims 1, 3 – 9, 11 – 15, 17 – 26 and 28 – 34 is **maintained**.

6. The Examiner notes that Applicant argues the Examiner's use of Official Notice. The Examiner submits Applicant's challenge is still inadequate for the same reasons discussed in the past Office Action. However, to advance prosecution, the Examiner provides Cichelli et al. to show evidence that a bit can be flipped based on a hash. See col 7 lines 12 – 30). The Official Notice remains but is now "evidenced by" Cichelli et al.

Claim Rejections - 35 USC § 101

7. 35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

8. Claims **1, 3 – 9, 12 – 15 and 17 – 19** are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter.

9. Claims **1, 3 – 9, 12 – 15 and 17 – 19** fail the machine-or-transformation test which is a two-branched inquiry. It may be shown that a process claim satisfies 35 USC § 101 by showing that a claim is tied to a particular machine or by showing that a claim transforms an article into a different state or thing. See Gottschalk v. Benson, 409 U.S. 63, 67 (1972). As to the first prong (machine), the Examiner cannot find any showing that these claims are attached to a specific machine. As to the second prong (transformation), the process claims do not transform a physical article into a different state or thing. The process claims are merely manipulating abstract data without regard to any physical article or object.

Claim Rejections - 35 USC § 102(a)

10. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (a) the invention was known or used by others in this country, or patented or described in a printed publication in this or a foreign country, before the invention thereof by the applicant for a patent.

11. Claims **13 and 17** are rejected under 35 U.S.C. 102(a) as being anticipated by Schleimer et al. ("Winnowing: Local Algorithms for Document Fingerprinting"; published 9 June 2003).

12. In regard to **claim 13**, Schleimer et al discloses sampling the document to obtain a plurality of overlapping samples (Section 3: Winnowing; windows can be overlapping sample of a document); generating a set of checksum values (figure 2(d)); choosing a set of checksum values (Section 3: Winnowing; figure 2(e) shows a set of

predetermined size of elements are selected into windows); and compacting the subset of the overlapping blocks to obtain the representation of the document (figure 2(e) shows the compaction by winnowing, so does Section 3); and setting bits in the fingerprint of the document ((g) of Figure 2, bits of the representation of the document i.e. document signature are set).

13. In regard to **claim 17**, Schleimer et al discloses hashing the overlapping blocks (figure 2(d)).

Claim Rejections - 35 USC § 103

14. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

15. Claims **1, 5, 6, 8, 11, 20, 24, 26 and 28** are rejected under 35 U.S.C. 103(a) as being unpatentable over Schleimer et al. ("Winnowing: Local Algorithms for Document Fingerprinting"; published 9 June 2003).

16. In regard to **claim 1**, Schleimer et al discloses sampling the document to obtain a plurality of overlapping blocks (Section 3: Winnowing; windows can be overlapping sample of a document); generating a set of checksum values (figure 2(d)); choosing a set of checksum values (Section 3: Winnowing; figure 2(e) shows a set of predetermined size of elements are selected into windows); and compacting the subset

of the overlapping blocks to obtain the representation of the document (figure 2(e) shows the compaction by winnowing, so does Section 3).

Schleimer et al teaches the choosing of less then all of the k-gram hashes. Therefore, the difference between the claimed invention of claim 1 and the Schleimer et al. reference is that the claimed invention selects the subset of blocks and then hashes the selected blocks ("compacting") and the Schleimer et al. reference hashes all of the k-grams (selected blocks) and then selects a smaller subset (for example; four selected hashes in step Figure 1c) to compose the fingerprint. Thus, this is an obvious reordering of steps that would be recognized by a person of ordinary skill in the art because it leads to increased efficiency and a reduction of data processing (selecting fewer then all blocks to hash reduces the amount of processor time). It would have been obvious to a person of ordinary skill in the art at the time of invention to reorder the steps of Schleimer et al. because it leads to improved efficiency.

17. In regard to **claim 5**, Schleimer et al discloses initializing the fingerprint before flipping bits in the fingerprint (variables have to be defined before they can be used).

18. In regard to **claim 6**, Schleimer et al discloses choosing the smallest hash value in a window (Section 3).

19. In regard to **claim 8**, Schleimer et al discloses hashing the overlapping blocks (figure 2(d)).

20. In regard to **claim 11**, Schleimer et al discloses the overlapping blocks can be of the same length (Section 3).

21. In regard to **claim 20**, Schleimer et al discloses sampling the document to obtain a plurality of overlapping blocks (Section 3: Winnowing; windows can be overlapping sample of a document); generating a set of checksum values (figure 2(d)); choosing a set of checksum values (Section 3: Winnowing; figure 2(e) shows a set of predetermined size of elements are selected into windows); and compacting the subset of the overlapping blocks to obtain the representation of the document (figure 2(e) shows the compaction by winnowing, so does Section 3). Schleimer et al further discloses comparing the document representation against a query based on the representation of another document (Section 3.2).

Schleimer et al teaches the choosing of less then all of the k-gram hashes. Therefore, the difference between the claimed invention of claim 1 and the Schleimer et al. reference is that the claimed invention selects the subset of blocks and then hashes the selected blocks ("compacting") and the Schleimer et al. reference hashes all of the k-grams (selected blocks) and then selects a smaller subset (for example; four selected hashes in step Figure 1c) to compose the fingerprint. Thus, this is an obvious reordering of steps that would be recognized by a person of ordinary skill in the art because it leads to increased efficiency and a reduction of data processing (selecting fewer then all blocks to hash reduces the amount of processor time). It would have been obvious to a person of ordinary skill in the art at the time of invention to reorder the steps of Schleimer et al. because it leads to improved efficiency.

22. In regard to **claim 24**, Schleimer et al discloses choosing the smallest hash value in a window (Section 3).

23. In regard to **claim 26**, Schleimer et al discloses sampling the document to obtain a plurality of overlapping blocks (Section 3: Winnowing; windows can be overlapping sample of a document); choosing a subset of the overlapping blocks (Section 3: Winnowing; figure 2(e) shows a set of predetermined size of elements are selected into windows); and compacting the subset of the overlapping blocks to obtain the representation of the document (figure 2(e) shows the compaction by winnowing, so does Section 3).

Schleimer et al teaches the choosing of less then all of the k-gram hashes. Therefore, the difference between the claimed invention of claim 1 and the Schleimer et al. reference is that the claimed invention selects the subset of blocks and then hashes the selected blocks ("compacting") and the Schleimer et al. reference hashes all of the k-grams (selected blocks) and then selects a smaller subset (for example; four selected hashes in step Figure 1c) to compose the fingerprint. Thus, this is an obvious reordering of steps that would be recognized by a person of ordinary skill in the art because it leads to increased efficiency and a reduction of data processing (selecting fewer then all blocks to hash reduces the amount of processor time). It would have been obvious to a person of ordinary skill in the art at the time of invention to reorder the steps of Schleimer et al. because it leads to improved efficiency.

24. In regard to **claim 28**, Schleimer et al discloses choosing the smallest hash value in a window (Section 3).

25. Claims **1 and 5 – 8** are rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) in view of Ward et al (US PGPUB 2002/0133499 A1; published 19 September 2002) and Schleimer et al. ("Winnowing: Local Algorithms for Document Fingerprinting"; published 9 June 2003) and .

26. In regard to **claim 1**, Burrows teaches sampling the document to obtain a plurality of blocks (see figure 4, blocks are selected from a document); generating a set of checksum values (figure 2(d)); choosing a set of checksum values (see figure 4); and compacting the subset of the overlapping blocks to obtain the representation of the document (Figure 5 shows the resultant compaction of the results of the selected overlapping blocks); setting bits in the representation of the document based on the subset of the overlapping blocks (Figure 5 shows the resultant compaction of the results of the selected overlapping blocks; since the representation is stored in memory or on a disk, bits are set based on the overlapping blocks).. Burrows does not explicitly teach that the blocks can be overlap data. However, Ward et al teaches a sliding window of overlap for data (see ¶32). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

Schleimer et al teaches the choosing of less then all of the k-gram hashes. Therefore, the difference between the claimed invention of claim 1 and the Schleimer et al. reference is that the claimed invention selects the subset of blocks and then hashes

the selected blocks ("compacting") and the Schleimer et al. reference hashes all of the k-grams (selected blocks) and then selects a smaller subset (for example; four selected hashes in step Figure 1c) to compose the fingerprint. Thus, this is an obvious reordering of steps that would be recognized by a person of ordinary skill in the art because it leads to increased efficiency and a reduction of data processing (selecting fewer then all blocks to hash reduces the amount of processor time). It would have been obvious to a person of ordinary skill in the art at the time of invention to reorder the steps of Schleimer et al. because it leads to improved efficiency.

27. In regard to **claim 5**, Ward et al teaches hashing of the data blocks (see ¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

28. In regard to **claim 6**, Ward et al teaches choosing the highest weighted feature of the computed vectors. Another obvious choice inferred from Ward et al is the lowest weighted feature can be chosen (¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

29. In regard to **claim 7**, Ward et al teaches choosing the highest weighted feature of the computed vectors (¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of

Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

30. In regard to **claim 8**, Ward et al teaches hashing of the data blocks (see ¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

31. Claims **3, 4, 11 and 12** are rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows, Ward et al and Schleimer et al. as applied to claim 1 above, and further in view of Broder et al (US Patent 6,230,155 B1; patented 8 May 2001).

32. In regard to **claim 3**, Burrows and Ward et al teach the invention as substantially claimed. Burrows and Ward et al do not explicitly teach that the representation of the document be of a predetermined length. However, Broder et al does teach that a predetermined length of the representation of a document (see col 5, lines 1 – 14). It would have been obvious to a person of ordinary skill in the art at the time of invention to use the predetermined length of representation of a document of Broder et al with the method of Burrows and Ward et al because it allows for easy comparison of the fingerprints between two documents.

33. In regard to **claim 4**, Burrows and Ward et al teach the invention as substantially claimed. However, Broder et al does teach that a predetermined length of the representation of a document (see col 5, lines 1 – 14) and suggests that a longer fingerprint reduces the chance of two documents that are not similar that have exactly

the same fingerprint (see col 5, lines 1 – 14). Therefore, it would have been obvious to a person of ordinary skill in the art at the time of invention to use the predetermined length of representation of a document of Broder et al with the method of Burrows and Ward et al because it allows for easy comparison of the fingerprints between two documents as well as reduce the chance of two documents that are not similar that have exactly the same fingerprint.

34. In regard to **claim 11**, Burrows and Ward et al teach the invention as substantially claimed. However, Broder et al teaches that "words" can be of a predetermined size such as 8 bytes (see col 6, lines 4 – 7). It would have been obvious to a person of ordinary skill in the art at the time of invention to use the predetermined length of representation of a document of Broder et al with the method of Burrows and Ward et al because it allows for easy comparison of the fingerprints between two documents as well as reduce the chance of two documents that are not similar that have exactly the same fingerprint.

35. In regard to **claim 12**, Burrows and Ward et al teach the invention as substantially claimed. However, Broder et al teaches that "words" can be of a predetermined size such as 8 bytes (see col 6, lines 4 – 7) and under sized words can be padded to bring them to correct size. It would have been obvious to a person of ordinary skill in the art at the time of invention to use the predetermined length of representation of a document of Broder et al with the method of Burrows and Ward et al because it allows for easy comparison of the fingerprints between two documents as

well as reduce the chance of two documents that are not similar that have exactly the same fingerprint.

36. Claims **13 and 17** are rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) in view of Ward et al (US PGPUB 2002/0133499 A1; published 19 September 2002).

37. In regard to **claim 13**, Burrows teaches sampling the document to obtain a plurality of blocks (see figure 4, blocks are selected from a document); choosing a subset of the overlapping blocks (see figure 4, blocks are selected from a document); setting bits in the representation of the document based on the subset of the overlapping blocks (Figure 5 shows the resultant compaction of the results of the selected overlapping blocks; since the representation is stored in memory or on a disk, bits are set based on the overlapping blocks). Burrows does not explicitly teach that the blocks can be overlap data. However, Ward et al teaches a sliding window of overlap for data (see ¶32). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

38. In regard to **claim 17**, Ward et al teaches hashing of the data blocks (see ¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

39. Claims **14 and 15** are rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows, Ward et al as applied to claim 13 above, and further in view of Broder et al (US Patent 6,230,155 B1; patented 8 May 2001).

40. In regard to **claim 14**, Burrows and Ward et al teach the invention as substantially claimed. Burrows and Ward et al do not explicitly teach that the representation of the document be of a predetermined length. However, Broder et al does teach that a predetermined length of the representation of a document (see col 5, lines 1 – 14). It would have been obvious to a person of ordinary skill in the art at the time of invention to use the predetermined length of representation of a document of Broder et al with the method of Burrows and Ward et al because it allows for easy comparison of the fingerprints between two documents.

41. In regard to **claim 15**, Burrows and Ward et al teach the invention as substantially claimed. However, Broder et al does teach that a predetermined length of the representation of a document (see col 5, lines 1 – 14) and suggests that a longer fingerprint reduces the chance of two documents that are not similar that have exactly the same fingerprint (see col 5, lines 1 – 14). Therefore, it would have been obvious to a person of ordinary skill in the art at the time of invention to use the predetermined length of representation of a document of Broder et al with the method of Burrows and Ward et al because it allows for easy comparison of the fingerprints between two documents as well as reduce the chance of two documents that are not similar that have exactly the same fingerprint.

42. Claim **20** is rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) in view of Schleimer et al. ("Winnowing: Local Algorithms for Document Fingerprinting"; published 9 June 2003).

43. In regard to **claim 20**, Burrows discloses a fingerprint creation unit (2410) that samples the document to obtain a plurality of blocks (see figure 4, blocks are selected from a document); chooses a subset of the overlapping blocks (see figure 4, blocks are selected from a document); and compacting the subset of the overlapping blocks to obtain the representation of the document (Figure 5 shows the resultant compaction of the results of the selected overlapping blocks). Burrows further discloses a similarity detection component to compare fingerprints to determine whether pairs of fingerprints correspond to near-duplicate documents (2420 and see figure 24).

However, the Examiner notes that Burrows does not appear to teach the limitation "where the subset is less then the entirety of the plurality of overlapping blocks". The Schleimer et al. reference selects a smaller subset (for example; four selected hashes in step Figure 1c) to compose the fingerprint. It would have been obvious to a person of ordinary skill in the art at the time of invention to use the selection capabilities of Schleimer et al. with the system/method of Burrows because it leads to improved efficiency.

44. Claims **22 and 23** are rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows and Schleimer et al. as applied to claim 20 above, and further in view of

Charikar ("Similarity Estimation Techniques from Rounding Algorithms"; published 19 May 2002).

45. In regard to **claim 22**, Burrows teaches the invention as substantially claimed. Burrows does not explicitly state the use of Hamming Space for the comparison of fingerprints. However, Charikar does teach use of Hamming space for calculating the similarity between fingerprints (see page 382, col 2, second paragraph). It would have been obvious to a person ordinary skill in the art at the time of invention to use the Hamming space calculations of Charikar with the components of Burrows because it is an able and suggested method for computing nearest neighbor problems and similarity tests.

46. In regard to **claim 23**, Burrows teaches the invention as substantially claimed. Burrows does not explicitly state the use of Hamming Space for the comparison of fingerprints. However, Charikar does teach use of Hamming space for calculating the similarity between fingerprints (see page 382, col 2, second paragraph). It would have been obvious to a person ordinary skill in the art at the time of invention to use the Hamming space calculations of Charikar with the components of Burrows because it is an able and suggested method for computing nearest neighbor problems and similarity tests.

47. Claims **24 and 25** are rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) and Schleimer et al. as

applied to claim 20 above, and further in view of Ward et al (US PGPUB 2002/0133499 A1; published 19 September 2002).

48. In regard to **claim 24**, Burrows teaches the invention as substantially claimed. Ward et al teaches choosing the highest weighted feature of the computed vectors. Another obvious choice inferred from Ward et al is the lowest weighted feature can be chosen (¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

49. In regard to **claim 25**, Burrows teaches the invention as substantially claimed. Ward et al teaches choosing the highest weighted feature of the computed vectors (¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

50. Claims **26, 28 and 29** are rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) and Schleimer et al. in view of Ward et al (US PGPUB 2002/0133499 A1; published 19 September 2002).

51. In regard to **claim 26**, Burrows teaches sampling the document to obtain a plurality of blocks (see figure 4, blocks are selected from a document); choosing a subset of the overlapping blocks (see figure 4, blocks are selected from a document); and compacting the subset of the overlapping blocks to obtain the representation of the document (Figure 5 shows the resultant compaction of the results of the selected

overlapping blocks). Burrows does not explicitly teach that the blocks can be overlap data. However, Ward et al teaches a sliding window of overlap for data (see ¶32). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

Schleimer et al teaches the choosing of less then all of the k-gram hashes. Therefore, the difference between the claimed invention of claim 1 and the Schleimer et al. reference is that the claimed invention selects the subset of blocks and then hashes the selected blocks ("compacting") and the Schleimer et al. reference hashes all of the k-grams (selected blocks) and then selects a smaller subset (for example; four selected hashes in step Figure 1c) to compose the fingerprint. Thus, this is an obvious reordering of steps that would be recognized by a person of ordinary skill in the art because it leads to increased efficiency and a reduction of data processing (selecting fewer then all blocks to hash reduces the amount of processor time). It would have been obvious to a person of ordinary skill in the art at the time of invention to reorder the steps of Schleimer et al. because it leads to improved efficiency.

52. In regard to **claim 28**, Burrows teaches the invention as substantially claimed. Ward et al teaches choosing the highest weighted feature of the computed vectors. Another obvious choice inferred from Ward et al is the lowest weighted feature can be chosen (¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

53. In regard to **claim 29**, Burrows teaches the invention as substantially claimed. Ward et al teaches choosing the highest weighted feature of the computed vectors (¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

54. Claim **31, 32 and 34 – 36** are rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) in view of Ward et al (US PGPUB 2002/0133499 A1; published 19 September 2002) and Schleimer et al. ("Winnowing: Local Algorithms for Document Fingerprinting"; published 9 June 2003).

55. In regard to **claim 31**, Burrows teaches sampling the document to obtain a plurality of blocks (see figure 4, blocks are selected from a document); choosing a subset of the overlapping blocks (see figure 4, blocks are selected from a document); setting bits in the representation of the document based on the subset of the overlapping blocks (Figure 5 shows the resultant compaction of the results of the selected overlapping blocks; since the representation is stored in memory or on a disk, bits are set based on the overlapping blocks). Burrows does not explicitly teach that the blocks can be overlap data. However, Ward et al teaches a sliding window of overlap for data (see ¶32). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

Schleimer et al teaches the choosing of less then all of the k-gram hashes. Therefore, the difference between the claimed invention of claim 1 and the Schleimer et al. reference is that the claimed invention selects the subset of blocks and then hashes the selected blocks ("compacting") and the Schleimer et al. reference hashes all of the k-grams (selected blocks) and then selects a smaller subset (for example; four selected hashes in step Figure 1c) to compose the fingerprint. Thus, this is an obvious reordering of steps that would be recognized by a person of ordinary skill in the art because it leads to increased efficiency and a reduction of data processing (selecting fewer then all blocks to hash reduces the amount of processor time). It would have been obvious to a person of ordinary skill in the art at the time of invention to reorder the steps of Schleimer et al. because it leads to improved efficiency.

56. In regard to **claim 32**, Ward et al teaches hashing of the data blocks (see ¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

57. In regard to **claim 34**, Burrows teaches sampling the document to obtain a plurality of blocks (see figure 4, blocks are selected from a document); choosing a subset of the overlapping blocks (see figure 4, blocks are selected from a document); setting bits in the representation of the document based on the subset of the overlapping blocks (Figure 5 shows the resultant compaction of the results of the selected overlapping blocks; since the representation is stored in memory or on a disk, bits are set based on the overlapping blocks). Burrows does not explicitly teach that the

blocks can be overlap data. However, Ward et al teaches a sliding window of overlap for data (see ¶32). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

Schleimer et al teaches the choosing of less then all of the k-gram hashes. Therefore, the difference between the claimed invention of claim 1 and the Schleimer et al. reference is that the claimed invention selects the subset of blocks and then hashes the selected blocks ("compacting") and the Schleimer et al. reference hashes all of the k-grams (selected blocks) and then selects a smaller subset (for example; four selected hashes in step Figure 1c) to compose the fingerprint. Thus, this is an obvious reordering of steps that would be recognized by a person of ordinary skill in the art because it leads to increased efficiency and a reduction of data processing (selecting fewer then all blocks to hash reduces the amount of processor time). It would have been obvious to a person of ordinary skill in the art at the time of invention to reorder the steps of Schleimer et al. because it leads to improved efficiency.

58. In regard to **claim 35**, Schleimer et al teaches initializing the fingerprint before flipping bits in the fingerprint (variables have to be defined before they can be used).

59. In regard to **claim 36**, Schleimer et al further discloses comparing the document representation against a query based on the representation of another document (Section 3.2).

60. Claim **9** is rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) in view of Ward et al (US PGPUB 2002/0133499 A1; published 19 September 2002) and Schleimer et al. as applied to claim 8 above, and further in view of Official Notice.

61. In regard to **claim 9**, Burrows and Ward et al teach the invention as substantially claimed. The Examiner takes Official Notice that taking a number of least significant bits is well known by a person of ordinary skill in the art at the time of invention. It would have been obvious to a person of ordinary skill in the art at the time of invention to use this type of hashing technique in the method of Burrows and Ward et al because it is allows a convenient low overhead method a determining which bin a particular sample gets placed into.

62. Claim **18** is rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) in view of Ward et al (US PGPUB 2002/0133499 A1; published 19 September 2002) and Schleimer et al. as applied to claim 17 above, and further in view of Official Notice.

63. In regard to **claim 18**, Burrows and Ward et al teach the invention as substantially claimed. The Examiner takes Official Notice that taking a number of least significant bits is well known by a person of ordinary skill in the art at the time of invention. It would have been obvious to a person of ordinary skill in the art at the time of invention to use this type of hashing technique in the method of Burrows and Ward et

al because it is allows a convenient low overhead method a determining which bin a particular sample gets placed into.

64. Claim **19** is rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) in view of Ward et al (US PGPUB 2002/0133499 A1; published 19 September 2002) as applied to claim 17, and further in view of Official Notice.

In regard to **claim 19**, Burrows and Ward et al teach the invention as substantially claimed. The Examiner takes Official Notice that flipping bits based on a hash (as it done for generation of encryption keys via hashing) is well known by a person of ordinary skill in the art at the time of invention. It would have been obvious to a person of ordinary skill in the art at the time of invention to use this type of hashing technique in the method of Burrows and Ward et al because it is allows a convenient low overhead method a determining which bin a particular sample gets placed into.

65. Claim **33** is rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) in view of Ward et al (US PGPUB 2002/0133499 A1; published 19 September 2002) and Schleimer et al. as applied to claim 32 above, and further in view of Official Notice.

66. In regard to **claim 33**, Burrows and Ward et al teach the invention as substantially claimed. The Examiner takes Official Notice that taking a number of least significant bits is well known by a person of ordinary skill in the art at the time of

invention. It would have been obvious to a person of ordinary skill in the art at the time of invention to use this type of hashing technique in the method of Burrows and Ward et al because it is allows a convenient low overhead method a determining which bin a particular sample gets placed into.

67. Claim **21** is rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) as applied to claim 20 above, and further in view of Official Notice.

68. In regard to **claim 21**, Burrows also discloses a search engine (140, see figure 1). However, Burrows does not explicitly disclose returning a single link when the documents are determined to be duplicates. The Examiner takes Office Notice that returning a single link when the documents are determined to be duplicates is well known by a person of ordinary skill in the art at the time of invention. It would have been obvious to a person of ordinary skill in the art at the time of invention to use returning a single link with the components of Burrows because it would reduce the amount of data traffic and provide the user with clarity as to the nature of the document.

69. Claim **30** is rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) in view of Ward et al (US PGPUB 2002/0133499 A1; published 19 September 2002) and Schleimer et al. as applied to claim 26, and further in view of Official Notice.

70. In regard to **claim 30**, Burrows and Ward et al teach the invention as substantially claimed. The Examiner takes Official Notice that flipping bits based on a hash (as it done for generation of encryption keys via hashing) is well known by a person of ordinary skill in the art at the time of invention. It would have been obvious to a person of ordinary skill in the art at the time of invention to use this type of hashing technique in the method of Burrows and Ward et al because it is allows a convenient low overhead method a determining which bin a particular sample gets placed into.

Conclusion

71. The Examiner requests, in response to this Office action, that support be shown for language added to any original claims on amendment and any new claims. That is, indicate support for newly added claim language by specifically pointing to page(s) and line no(s) in the specification and/or drawing figure(s). This will assist the Examiner in prosecuting the application.

72. When responding to this Office action, Applicant is advised to clearly point out the patentable novelty which he or she thinks the claims present, in view of the state of the art disclosed by the references cited or the objections made. He or she must also show how the amendments avoid such references or objections See 37 CFR 1.111(c).

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Garrett Smith whose telephone number is (571)270-1764. If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Tim T. Vo can be reached on (571) 272-3642. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

February 3, 2009

/GS/
Garrett Smith
Patent Examiner
Art Unit 2168

/Tim T. Vo/

Supervisory Patent Examiner, Art Unit 2168